GEOPHYSICS, VOL. 56, NO. 4 (APRIL 1991); P. 472–482, 11 FIGS.

Bounding seismic velocities using a tomographic method

D. W. Vasco*

ABSTRACT

Traveltime data are used to determine upper and lower bounds on velocity variations in the earth by an iterative method. In effect, the range of models is found which is consistent with the data, rather than a single model which is "best fitting" in some sense. The algorithm used, a variant of the row-action algorithms commonly applied to tomographic inversions, requires little in-core memory and has proven feasible for data sets of the order of hundreds of thousands to a million traveltimes. Any inequality constraint, such as that the velocity is always positive, may be incorporated into the formulation. Data errors can be included both locally (as strict constraints on each datum) and globally (as a constraint on a total error measure). The method may also be used to derive the velocity structure which results in the minimum ℓ^1 norm of the residual misfit.

Data from the Grimsel crosshole experiment are used to map confidence bounds of 0.02 ms on 1521 traveltime residuals into upper and lower bounds on seismic velocities. There is great variation in the widths of these bounds as a function of subsurface position from 0.1 to 0.9 km/s. The distributions of the bounds agree with the parameter resolution values found from a singular value decomposition (SVD) and suggest that a low-velocity mylonitic zone, seen in tomographic inversions of the traveltime data, is adequately imaged. Though the data were corrected for seismic anisotropy, significant alternating positive and negative velocity perturbations in poorly constrained quadrants of the crosshole region suggest that some residual anisotropy effects are still present.

INTRODUCTION

Tomographic inversions of seismic traveltime data are frequently used to determine subsurface velocities. The usefulness of this method derives from the fact that the observed traveltimes T are line integrals of the reciprocal of the velocity $c\ (x, y, z)$ along a raypath $R\ (x, y, z)$,

$$T = \int_{R(x,y,z)} \frac{dR}{c(x, y, z)}.$$

The traveltime is nonlinearly related to the velocity since the raypath geometry depends upon the velocity structure. The problem may be linearized by fixing the raypaths [$R\ (x, y, z) = R_0\ (x, y, z)$] using some initial velocity model $c_0\ (x, y, z)$ and solving for some parameterized perturbations in velocity δc (Nolet, 1985; Vasco, 1986),

$$\delta T = \int_{R_0(x,y,z)} \frac{\delta c}{c_0(x, y, z)^2}\, dR.$$

In a tomographic experiment many sources transmit seismic energy to receivers distributed throughout a region. For each source-receiver pair, a linear equation relates the velocity perturbation parameters to the observed traveltime. The total set of observations determines a system of equations which may be solved for velocity perturbations, expressed in matrix-vector notation as

$$\mathbf{t} = \underline{\mathbf{G}}\mathbf{v}. \tag{1}$$

The M by N matrix $\underline{\mathbf{G}}$ contains the influence of N parameters describing the velocity perturbations on the M traveltime observations. The vector on the left-hand side contains the traveltime residuals $t_i = \delta T_i = T_i^{\text{Observed}} - T_i^{\text{Calculated}}$; the vector on the right contains the velocity perturbations from

Manuscript received by the Editor February 23, 1990; revised manuscript received July 27, 1990.
*Department of Geology and Geophysics, University of California, Berkeley, CA 94720; also at Center for Computational Seismology, Earth Sciences Division/Building 50E, Lawrence Berkeley Laboratory, 1 Cyclotron Road, Berkeley, CA 94720.

the average structure $v_j = \delta c_j$. This is often a very large system of equations, with M of the order of 10^6. Commonly, tomographic problems are even large enough to preclude storage of all the elements of $\mathbf{G}$ and $\mathbf{t}$ in core memory. Two frequently used approaches for solving large linear systems of equations are a conjugate gradient-based algorithm (LSQR) (Paige and Saunders, 1982; Nolet, 1985; Scales 1987) and row-action methods (ART, SIRT) (Herman and Lent, 1978a,b; Dines and Lytle, 1979; Lent and Censor, 1980). Both formulations are iterative and row-active, i.e., they only operate on a single row of equations for a given iteration. Both the SIRT and LSQR algorithms have enabled seismologists to invert data sets on the order of millions of traveltimes for thousands to tens of thousands of unknown velocity parameters (Clayton and Comer, 1983; Spakman and Nolet, 1988).

Once a solution to the inverse problem has been computed, it is necessary to assess the resulting velocity model. That is, how well the data determine the model parameters should be quantified. In addition, some estimate of the errors associated with the derived model parameters is desirable. Conventionally, this is accomplished through the formation of the generalized inverse (Aki and Richards, 1980; Menke, 1984) from which the parameter resolution matrix and the parameter covariance matrix are derived. The parameter resolution matrix $\mathbf{R}$ relates the estimated subsurface structure $\hat{\mathbf{v}}$ to the "true" velocity structure, $\mathbf{v}_t$,

$$\hat{\mathbf{v}} = \mathbf{R}\mathbf{v}_t.$$

The rows of $\mathbf{R}$ are averaging coefficients of the "true" subsurface velocity. When the problem is completely underdetermined, the resolution matrix of the generalized inverse is written

$$\mathbf{R} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}.$$

The superscript T signifies a matrix transpose. If $\mathbf{R}$ were the identity matrix, the subsurface velocity structure would be perfectly resolved. The parameter covariance matrix $\mathbf{C}_{vv}$ is the result of mapping the errors in the data to the errors in the model parameters (Menke, 1984). For an underdetermined problem with statistically independent data with variance σ^2, the estimated errors in the model parameters are given by

$$\mathbf{C}_{vv} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-2}\mathbf{G}\sigma^2$$

for the generalized inverse (Aki and Richards, 1980; Menke, 1984).

For the large systems of equations associated with tomographic problems, calculation of these matrices is often not feasible. The number of operations (additions and multiplications) involved in estimating the generalized inverse is of order N^3 (Gill et al., 1981), where N is the number of unknown parameters characterizing the velocity structure. There is also a problem in that a tradeoff exists between the model resolution and model covariance matrices which is a function of the damping associated with the generalized inverse (Aki and Richards, 1980).

This paper takes another approach for assessing tomographic models. Specifically, I present a technique for mapping traveltime observations and their associated confidence bounds into upper and lower bounds on seismic velocities. There are a number of advantages to this approach. For a given discretization, data set, and error distribution, these bounds are unique. Also, this is a direct mapping from data and data confidence bounds to the range of subsurface velocities. The information contained in the bounds can be used, as would the model covariance matrix and the model resolution matrix, to assess a tomographic solution. In addition, the method may be used to evaluate the appropriateness of various possible discretizations for a given imaging problem. Furthermore, for a parameterization in terms of constant-velocity blocks, each bound constrains the velocity in a particular block, allowing one to focus on specific regions of interest. Finally, a significant advantage to this approach (discussed further in the following section) is that the bounds may be determined in a row-active manner, so that very large tomographic problems can be treated.

VELOCITY BOUNDS

The general technique for calculating velocity bounds, based on methods developed by Parker (1972, 1974, and 1975) and Sabatier (1977a,b), has been described in Vasco (1986) and will only be briefly discussed here. The basic idea is to minimize or maximize a specific velocity perturbation parameter, say v_j, e.g., the jth block when using a discretization in terms of constant velocity blocks. The possible velocity perturbations are constrained by the data and some confidence interval on the data,

$$\mathbf{t} - \boldsymbol{\sigma} \le \mathbf{G}\mathbf{v} \le \mathbf{t} + \boldsymbol{\sigma}. \tag{2}$$

The elements of the vector $\boldsymbol{\sigma}$ are the fixed confidence bounds on individual datums, i.e., these are point confidence intervals. Alternatively, it is possible to constrain the data using a statistical confidence bound,

$$\begin{aligned} &\mathbf{t} - \mathbf{e} \le \mathbf{G}\mathbf{v} \le \mathbf{t} + \mathbf{e} \\ &\sum_{i=1}^{M} e_i \le E \\ &\mathbf{e} \ge \mathbf{0}. \end{aligned} \tag{3}$$

e_i are nonnegative error variables and E is a bound on the total misfit. The elements of the vector $\mathbf{e}$ may be interpreted as the absolute values of the errors associated with the traveltime residuals. Hence, the value of E may be derived from the statistics of a sum of absolute values as presented by Parker and McNutt (1980). Specifically, they examined the statistical properties of the function

$$m = \sum_{i=1}^{M} \frac{|x_i|}{S_i},$$

where x_i are independent normal random variables with zero mean and standard error S_i. The quantity $|x_i|$ represents the nonnegative error variables e_i introduced above. Parker and McNutt found the mean $\bar{m}$ and variance var (m) of m,

$$\bar{m} = (2/\pi)^{1/2} N,$$

$$\text{var}\,(m) = (1 - 2/\pi) N,$$

as well as the probability that m is less than a given value E:

$$P(m \leq E) = (2/\pi)^{N/2} \int_D e^{-\mathbf{x}\cdot\mathbf{x}/2}\, d^N x.$$

The domain of integration D is defined by

$$x_i \geq 0$$

and

$$\sum_{i=1}^{N} x_i \leq E.$$

These formulas can be used to find the quantity E such that $P\,(m \leq E)$ is some specified value.

Maximizing or minimizing the velocity perturbation v_j subject to the above linear inequality constraints, equation (2) or equation (3), is termed a linear programming problem. The standard algorithm for solving such linear programming problems, the simplex algorithm (Dantzig, 1963; Murtagh and Saunders, 1978), is not suitable for really large tomographic problems. It does not operate in a row-active manner and requires a prohibitive core memory allocation for such problems. Thus, it suffers from the same limitations as the calculation of the resolution and covariance matrices. Furthermore, the number of mathematical operations required to solve a linear programming problem via the simplex algorithm is an exponential function of the number of unknowns, which grows rapidly with problem size. For these reasons I have developed the following row-action algorithm to solve the linear programs presented above. Because it only operates on a single row of G per iteration, large traveltime data sets may be considered.

ITERATIVE ALGORITHM FOR CONSTRUCTING VELOCITY BOUNDS

The iterative algorithm detailed below is based on two mathematical points: (a) under certain conditions the solution of a linear programming problem will be identical to the solution of a quadratic perturbation of the same problem (Mangasarian and Meyer, 1979; Mangasarian, 1981); (b) there exists an iterative procedure for solving the resulting quadratically perturbed problem which only requires the storage of one row of $\mathbf{G}$ per iteration (Hildreth, 1957; Herman and Lent, 1978a; Lent and Censor, 1980). The resulting algorithm is only slightly more complicated than row-action techniques presently used to solve tomographic problems. The necessary conditions, presented with the details of the algorithm, hold for the linearized traveltime problem.

Mangasarian and Meyer (1979) have proven that if the linear program

$$\min \mathbf{c}^T\mathbf{v} \tag{4}$$

subject to

$$\mathbf{G}\mathbf{v} \leq \mathbf{u}$$

has a solution, there exists a positive number $\bar{\varepsilon}$, such that for $0 \leq \varepsilon \leq \bar{\varepsilon}$ the solution of the quadratically perturbed problem

$$\min \frac{\varepsilon}{2} \mathbf{v}^T\mathbf{Q}\mathbf{v} + \mathbf{c}^T\mathbf{v} \tag{5}$$

subject to

$$\mathbf{G}\mathbf{v} \leq \mathbf{u}$$

is also the solution of the linear program. The bound $\mathbf{u}$ represents the other side of the inequalities in equation (2), i.e., $\mathbf{t} + \boldsymbol{\sigma}$ for the local constraints. For the statistical constraints [equation (3)], the generalizations of $\mathbf{G}$, $\mathbf{v}$, and $\mathbf{u}$ are straightforward. The value of $\bar{\varepsilon}$ is related to the Lagrange multipliers of the problem (Mangasarian and Meyer, 1979). For our linearized traveltime formulation [equation (1)] to hold, the velocity perturbations must be small, of the order of 10 or 20 percent, relative to the background velocity. Under such conditions the quadratic terms in the quantity to be minimized are small relative to the linear terms if the elements of $\mathbf{Q}$ and $\mathbf{c}$ are of order unity. For the computation of velocity bounds considered here, the elements are unity and so the solution of the quadratic perturbation will be an adequate approximation to the solution of the linear programming problem.

It remains to solve the perturbed quadratic problem for the bounds on velocity [equation (5)]. Lent and Censor (1980) have shown how to write the quadratic optimization problem in the form of a norm minimization problem. Once the optimization problem has been so transformed, a row-action algorithm may be devised to solve it. Consider the quadratically perturbed optimization problem

$$\min \frac{1}{2} \mathbf{v}^T\mathbf{Q}\mathbf{v} + \mathbf{c}^T\mathbf{v}$$

subject to

$$\mathbf{G}\mathbf{v} \leq \mathbf{u}$$

where $\mathbf{Q}$ is any positive definite matrix and $\mathbf{v}$ is a vector in R^n. In what follows, the elements of $\mathbf{Q}$ have been scaled by ε. The matrix $\mathbf{Q}$ has a Choleski decomposition $\mathbf{Q} = \mathbf{L}^T\mathbf{L}$, where $\mathbf{L}$ is a lower triangular matrix, and so the transformation

$$\mathbf{v} = \mathbf{L}^{-1}\mathbf{x} - \mathbf{Q}^{-1}\mathbf{c}$$

results in the following norm minimization form for the problem (Lent and Censor 1980),

$$\min \frac{1}{2} \mathbf{x}^T\mathbf{x}$$

subject to

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}$$

where $\mathbf{A} = \mathbf{G}\mathbf{L}^{-1}$, $\mathbf{b} = \mathbf{u} + \mathbf{G}\mathbf{Q}^{-1}\mathbf{c}$. This form is limiting for optimization of general quadratic functions because $\mathbf{Q}^{-1}$ and $\mathbf{L}^{-1}$ must be calculated but, as shall be seen, it is entirely adequate for our approximate linear programming problem.

Herman and Lent (1978a) presented an iterative row-action procedure for solving the norm minimization problem

subject to general inequality constraints and simple interval constraints on the variables

$$\min \mathbf{x}^T\mathbf{x}$$

subject to

$$\ell_i \le \mathbf{a}_i^T\mathbf{x} \le u_i, \quad i = 1, 2, \ldots, M$$

$$\lambda_j \le x_j \le \mu_j, \quad j = 1, 2, \ldots, N.$$

Here, $\mathbf{a}_i^T\mathbf{x}$ denotes the inner or dot product of the ith row of $\underline{\mathbf{A}}$ with the vector $\mathbf{x}$. Upper and lower bounds are given by u_i and ℓ_i for the general linear constraints and by μ_j and λ_j for the interval constraints (a priori upper and lower bounds) on the velocities. Herman and Lent's formulation, based on an algorithm proposed by Hildreth (1957), introduces no new variables and involves computationally simple operations. First, initialize the vector $\mathbf{x}$ and a subsidiary vector $\mathbf{z}$, whose elements are known as dual variables,

$$\mathbf{x}^0 = 0$$

$$\mathbf{z}^0 = 0.$$

Then update these vectors at each iteration, for the ith row and the kth update,

$$\mathbf{x}^{k+1} = \mathbf{x}^k + sc^k\mathbf{a}_i \tag{6}$$

$$\mathbf{z}^{k+1} = \mathbf{z}^k + sc^k\mathbf{e}_i,$$

where $\mathbf{e}_i$ is a vector in which the only nonzero element is a 1 in the ith position, c^k is the median value of three elements,

$$c^k = \text{mid}\begin{Bmatrix} u_i - \mathbf{a}_i^T\bar{\mathbf{x}}^k/\mathbf{a}_i^T\mathbf{a}_i \\ z_i^k \\ \ell_i - \mathbf{a}_i^T\bar{\mathbf{x}}^k/\mathbf{a}_i^T\mathbf{a}_i \end{Bmatrix},$$

and

$$\bar{x}_j^k = \begin{cases} \lambda_j & x_j^k \le \lambda_j \\ x_j^k & \lambda_j \le x_j^k \le \mu_j, \\ \mu_j & \mu_j \le x_j^k \end{cases} \quad j = 1, 2, \ldots, N.$$

The factor s is termed the relaxation parameter and lies within the range $0.5 \le s \le 2.0$. In some situations s can significantly enhance the rate of convergence of the algorithm. The convergence of this algorithm with the relaxation parameter equal to 1.0 has been proven by Herman and Lent (1978a). When $s \ne 1.0$, convergence has only been proven for the general inequalities of equation (5) (without interval constraints). Thus, for a range (possibly infinite) of ε, a row-action method can be used for linear programming. In this case the matrix $\mathbf{Q}$ is given by $\varepsilon\underline{\mathbf{I}}$, a scaled identity matrix, and so $\mathbf{Q}^{-1}$ and $\underline{\mathbf{L}}^{-1}$ are trivially determined.

The iterative algorithm presented above may also be used to construct a model that minimizes the ℓ^1 norm of the residuals (Gill et al., 1981), that is, to find the model $\mathbf{v}$ for which

$$\sum_{j=1}^{M} \left| \sum_{i=1}^{N} G_{ij}v_i - t_j \right|$$

is a minimum. Such a model is the solution of the following linear programming problem:

$$\min \sum_{i=1}^{M} (x_i + y_i)$$

subject to

$$\underline{\mathbf{G}}\mathbf{v} - \mathbf{x} + \mathbf{y} = \mathbf{t}, \tag{7}$$

$$x_i \ge 0, \; y_i \ge 0, \; i = 1, M,$$

where $\underline{\mathbf{G}}$, $\mathbf{v}$, and $\mathbf{t}$ are defined as in equation (1). The variables x_i and y_i are nonnegative quantities which may take any value. Such a solution is less affected by outliers in the data $\mathbf{t}$ than is the least-squares (the minimum ℓ^2 residual norm) solution (Claerbout and Muir, 1973), and such a result is termed a robust estimate (Press et al., 1986).

To compare results of the iterative procedure developed here with simplex method solutions, a synthetic data set was created (Figure 1). A linear trend of 0.5 km/s velocity perturbations (darkened squares) diagonally cross a constant-velocity background of 5.3 km/s. A large number of raypaths sample the region, though the ray coverage is highly nonuniform. The confidence bounds on the data were arbitrarily set at 10 percent of the maximum traveltime residual for this comparison. The resulting traveltime residuals and the set of confidence bounds were used to compute upper bounds on possible velocity perturbations. The data confidence bounds were applied as point intervals, in the form of the system of equations (2). Figure 2 contrasts solutions from the simplex and the iterative row-action methods. The high-velocity zone from the center left to the lower right is clearly visible in both solutions. In general, the iterative row-action method gives very similar results to the simplex approach. The maximum upper bound of 0.63 km/s from the iterative method is very close to the maximum of 0.64 km/s produced by the simplex algorithm. The small differences were due to incomplete convergence because the algorithm was stopped after 30 iterations. Thus, this synthetic test gives us confidence in the iterative algorithm.

THE GRIMSEL CROSSHOLE EXPERIMENT

Experiment description

The Grimsel Rock Laboratory (GRL), operated by the Swiss National Cooperative for the Storage of Radioactive Waste (NAGRA), studies the effects of fractures on nuclear waste storage. A cooperative experiment between NAGRA and the Department of Energy's Lawrence Berkeley Laboratory was devised to demonstrate the applicability of seismic tomographic methods for fracture detection (Majer et al., 1990). The experiment provided an excellent opportunity to test the above algorithms on a realistic problem. Furthermore, the data set was small enough to allow the computation of the covariance and resolution matrices for comparison with the velocity perturbation bounds.

The laboratory is situated within granitic rock containing a foliation, defined by aligned grains of biotite, which strikes northeast and dips about 65 degrees to the southeast (Majer et al., 1990). In addition, a linear fabric is defined by elongated feldspar grains which parallel the foliation, giving

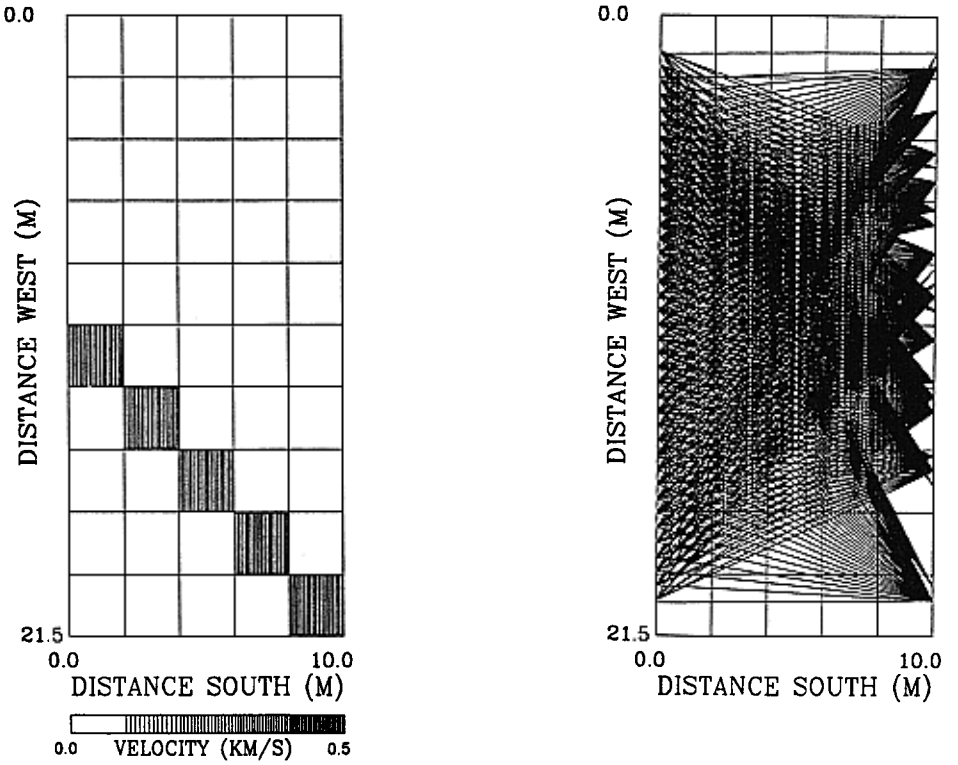

FIG. 1. (Left panel). Synthetic tomographic test problem. The discretization of the region into constant-velocity blocks is shown. The solid filled blocks contain positive velocity perturbations of 0.5 km/s. (Right panel). Raypath geometry and discretization for the synthetic tomographic problem.

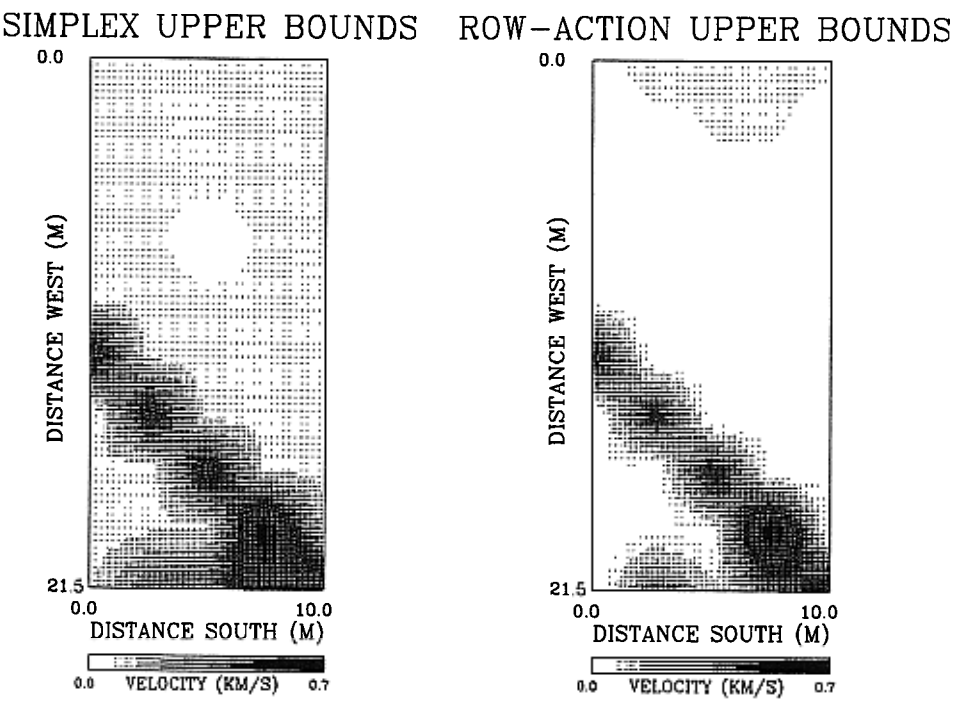

FIG. 2. (Left panel). Velocity upper bounds generated by the simplex algorithm. (Right panel). Upper bounds produced by the iterative row-active algorithm [equation (6)] for linear programming.

rise to a seismic anisotropy which has been estimated to be 10 percent (Majer et al., 1990). The study region is intersected by a 5 m thick mylonitic fracture zone (Figure 3) aligned with the foliation. From wall exposures and borehole information the overall structure of this shear zone, which consisted of a network of braided fractures, was known. The shear zone was expected to appear on any tomographic inversion of the Grimsel traveltime observations.

The experimental setup, shown in Figure 3, consisted of two north-south oriented tunnels which were connected via two subhorizontal boreholes (BOFR 87.001 and BOFR 87.002). The boreholes, 21.5 m long and separated by 10.0 m, contained *P*-wave sources spaced 0.5 m apart. Three-component accelerometers were also clamped within the boreholes, at the same interval, and their output digitized at 250 000 samples/s. To account for the known anisotropy, *P*-wave traveltimes were corrected by removing a low-order trigonometric function of ray azimuth from the arrival times (Backus, 1965). The correction removed a constant background anisotropy of nearly 10 percent. The resultant traveltimes had a strong linear trend as a function of source-receiver spacing, with a slope of 5.3 km/s (Figure 4). This slope was used to derive an initial background velocity to which velocity perturbations were referenced. The linearity of the time-distance plot suggests that an assumption of straight raypaths is valid. A histogram of traveltime residuals, computed relative to the constant background velocity, is shown in Figure 4. The mean of the residual distribution is 0.007 ms and the standard deviation is 0.026 ms. The distribution has clear non-Gaussian features including bimodality and no tail. The non-Gaussian features observable in Figure 4 are likely due to structural factors such as deviations from the initial velocity model.

The peak transmitted energy was 5000 to 10 000 Hz, resulting in approximate wavelengths of about 1.0 to 0.5 m

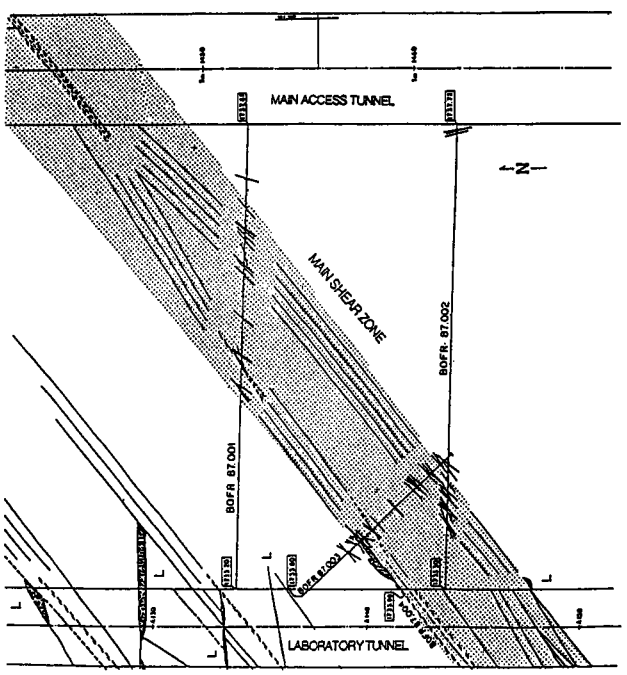

FIG. 3. Experimental setup for the Grimsel crosshole experiment (from Majer et al., 1990). The study region is bounded by two boreholes (BOFR 87.001 and BOFR 87.002) as well as by two access tunnels on the east and west edges. A mylonite shear zone is indicated by shading. The solid lines within the shear zone are hypothesized fracture orientations based on borehole measurements and extrapolations. 1.0 cm = 3.6 m.

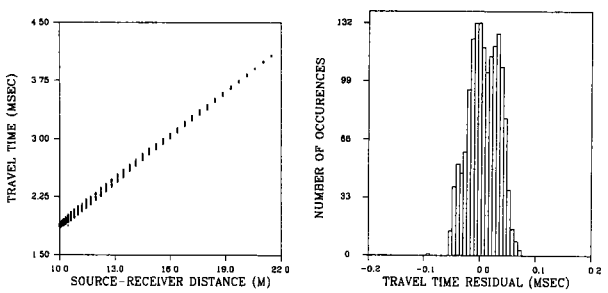

FIG. 4. (Left panel). Observed traveltimes versus source-receiver distance. All 1521 traveltime values are plotted as points. (Right panel). Traveltime residual histogram.

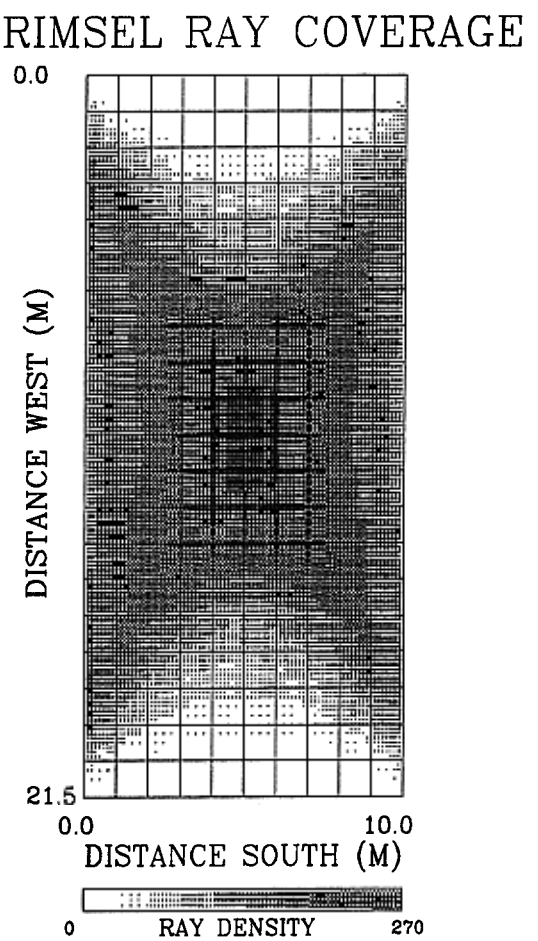

FIG. 5. Grimsel ray coverage. The discretization of the region used for the inversion (10 blocks along the north-south direction, 20 blocks along the east-west direction) may also be seen in this figure.

given the background velocity. The region in Figure 3 was discretized into a 10 (north-south) by 20 (east-west) grid with respective block dimensions of 1.000 by 1.075 m, the same order as the predominance wavelength. The total number of source-receiver pairs used in this study was 1521, producing excellent ray coverage (Figure 5) except at the extreme eastern and western edges. The ray coverage is nonuniform with many rays intersecting the center of the rectangular region, away from the four corners. Ray densities of up to 270 rays per block are achieved in this area. The blocks with the poorest coverage form rough V-shaped areas emanating from the eastern and western ends of the site. Constraints on block velocity often depend upon other factors besides ray density, such as the angular span of the rays through a given block. In addition, if a particular block is sampled only by raypaths traveling through many other blocks, tradeoffs in block velocities can occur.

SVD, LSQR, and ℓ^1 solutions

Because the data set was small enough to compute the generalized inverse, the resolution and covariance matrices could be calculated for this problem. This was accomplished using the singular value decomposition (SVD) of the matrix $\mathbf{G}$ in equation (1) (Aki and Richards, 1980; Menke, 1984; Press et al., 1986) to form the generalized inverse. Singular values less than 1/20th of the maximum singular value were set to zero, a rather stringent cutoff, to achieve adequate parameter resolution and satisfactory parameter errors. The resolution of the block velocities is fairly high (Figure 6) and when the full matrix was examined there appeared to be little tradeoff between blocks. The poorest resolution is found at the easternmost and westernmost edges of the GRL area. The parameter covariances (Figure 6) are more uniformly distributed, with all standard errors equal to or less than 0.1 km/s. The very low errors at the extreme east and west ends of the region occur because no rays traverse the blocks there. Therefore, they are not considered in calculating the resolution and covariances and the values are set to zero.

The generalized inverse was used to construct a solution incorporating the same singular value cutoff (Figure 7). The southeast corner is a region of high velocities, as is the northwest corner. Two low-velocity zones are seen extending from the northeast and north-central regions to the southwest. The velocity decrease of these zones is greater than -0.5 km/s, about five times greater than the errors found there. Thus, the mylonitic shear zone in this image is significant though somewhat obscured by the other features.

To see the variability possible in different tomographic solutions, consider the LSQR inversion and the solution which minimizes the ℓ^1 norm of the residuals (Figure 8). In the LSQR result, low velocities form a zone from the north center to the southwest, probably representing the mylonitic shear zone. The high-velocity zone in the southeast, seen in the SVD model, is visible in this solution also, evidence that it is not an artifact of the inversion algorithm. The maximum and minimum velocities for the LSQR solution were located on the eastern and western edges of the model, unlike the SVD solution. This is likely due to the stringent singular value cutoff used in the SVD algorithm damping out these large, poorly determined velocity perturbations. The minimization of the ℓ^1 norm of the residuals is accomplished using the row-action linear programming algorithm to solve

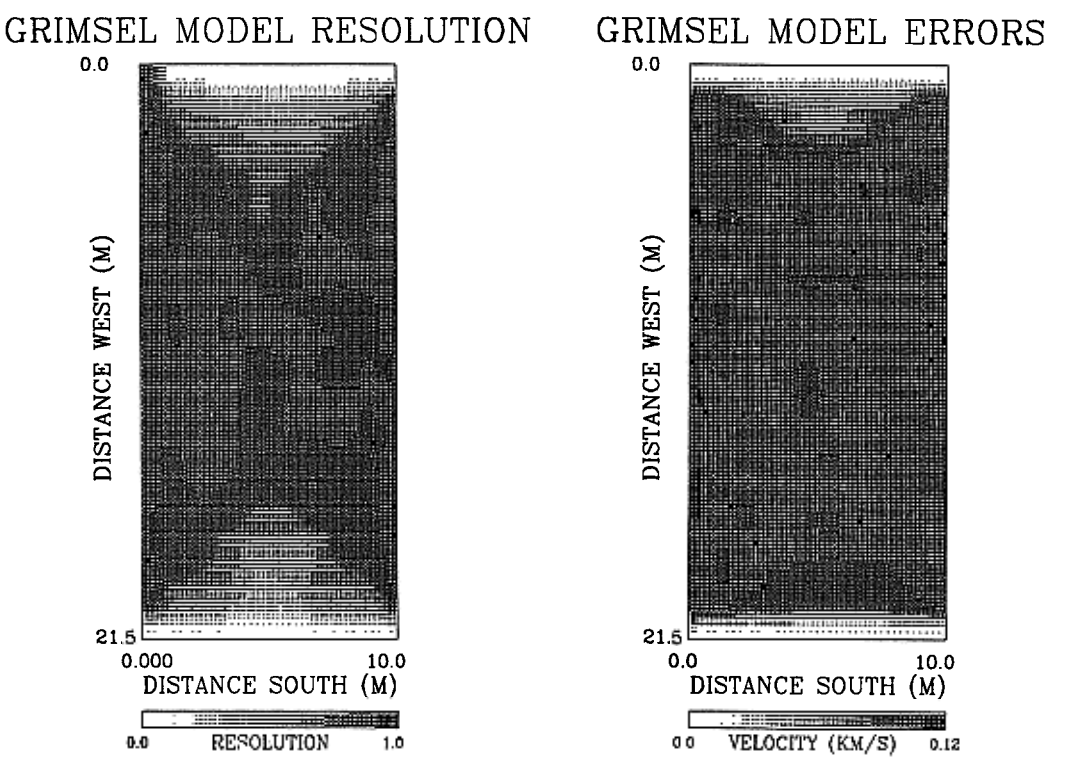

FIG. 6. (Left panel). Diagonal elements of the resolution matrix displayed in their corresponding geographic position. (Right panel). Square root of the diagonal elements of the covariance matrix, an estimate of the standard error.

the system of equalities and inequalities described earlier [equation (7)]. It took no more time to find this solution than to compute the LSQR solution. One can see similar features between this velocity model and the SVD-derived model. In particular, a zone of low velocities in the central region is discernible, as are the areas of high velocity in the southeast and northwest. The features are also coincident with those of the LSQR velocity model. Absent in the SVD model are the low-velocity perturbations present in the northeast and southwest quadrants of both the LSQR and the minimum ℓ^1 residual norm solutions. In addition, there are differences in the continuity of the two linear low-velocity features extending from the northeast and north central edges of the region to the south central and southwestern edges, respectively. Thus, the three models LSQR, SVD, and ℓ^1 contain some similarities as well as differences.

Velocity bounds

Given the differences in the above solutions, definitive statements can still be made about solutions satisfying the traveltime data. The iterative row-action method described earlier was used to construct statistically based upper and lower bounds on the velocity perturbations, i.e., the maximum and minimum possible perturbation in each block, subject to the constraints that the traveltime observations are satisfied within a specified confidence bound. The value of σ, one-half of the confidence interval, was chosen by an analysis of the traveltime residuals. The standard errors associated with the traveltime residuals from the LSQR model (Figure 9) were used as estimates of the random errors of the traveltimes. Implicit in this approach is the assumption that the LSQR model is able to remove all velocity structure effects from the residuals. That this is not exactly so can be seen in the histogram of Figure 9. When the velocity perturbations from the LSQR solution were added to the background velocity and used to compute the traveltime residuals, the residual distribution shown in Figure 9 resulted. Some of the features in the original residuals (Figure 4) are no longer present here. Slight tails have been introduced and the distribution is now skewed very slightly to the left due to the fact that straight raypaths (nonminimum time paths) were used to compute the residuals. The standard error associated with the distribution of residuals has been reduced to 0.02 ms.

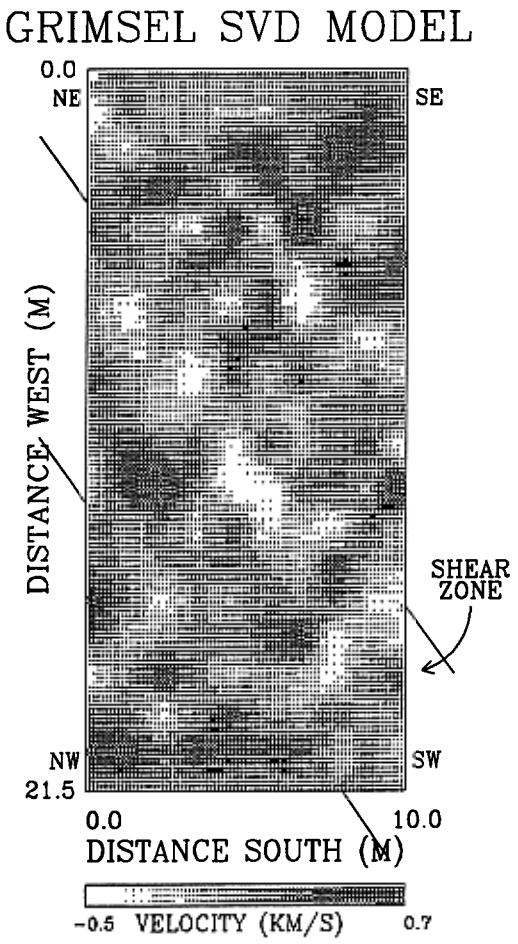

FIG. 7. Generalized inverse, computed using an SVD decomposition. The orientation and boundaries of the mylonite shear zone extend from the plot as diagonal lines.

The statistical confidence intervals of equation (3) were used to constrain the range of possible models. That is, each data point is not required to lie within its respective bounds. Instead, it is the total misfit of the traveltime data which is bounded. One standard error, as used in the formulas presented in Parker and McNutt (1980), was used to derive a value for E, the statistical error bound. The probability that the given data lie in the interval was set to 0.68 by specifying the appropriate value of E. A priori bounds on the velocity perturbations were applied

$$-1.5 \text{ km/s} \leq v \leq 2.0 \text{ km/s}$$

as additional inequality constraints in equation (3). These were arrived at by considering the maximum possible velocity perturbations in the three derived solutions (Figures 7 and 8). Such limits proved useful in controlling tradeoffs between blocks on a given raypath. Under these constraints, the lower bound on the velocity in each block was successively computed. The lowest velocity bounds are found in the eastern and western portions of the study area. The high-velocity perturbations in the southeast and northwest also appear in the bounds. The upper bounds were also derived, and there was a wide range in values from -0.7 km/s to 1.1 km/s. Regions of large positive velocity bounds are found in the northwest and southeast as in the lower bounds, SVD, and LSQR solutions. Significantly, blocks in the north-central and south-central areas are required to contain negative velocities, consistent with the derived velocity models, which contained low-velocity zones in those regions.

The mean value of the bounds, one-half of the upper plus the lower velocity bounds (Figure 10), may be compared to the solutions of the tomographic problem using the SVD technique (Figure 7), the LSQR method, and the minimum ℓ^1 residual norm algorithm (Figure 8). Again, alternating positive and negative velocity perturbations are found in the corners of the model. Only vague suggestions of the two low-velocity zones, linearly trending southwest, are visible in this figure. The half-width of the bounds, i.e., one-half the

upper bounds minus the lower bounds, should correlate somewhat with the standard errors and resolution of the model parameters. To check this, I computed one-half the difference between the computed upper and lower bounds which is shown in the right panel of Figure 10. In general there is good agreement between the bound half-widths and the parameter resolution shown in Figure 6. The widest possible ranges in velocity perturbation are found in the eastern and western edges, the maximum values of over 0.9 km/s. The half-width of the bounds, in the region away from the eastern and western edges, ranged from 0.1 km/s to 0.2 km/s which was slightly larger than, but of the order of, the standard errors computed from the SVD-based covariance estimates (0.06 km/s to 0.09 km/s). The differences between the two are due to the singular value cutoff used in the SVD decomposition as well as the value of E taken in equation (3). There is also only a rough correlation between the half-width of the bounds on the velocities and the ray-density distribution shown in Figure 5. The bounds are not as symmetric as the ray densities due to the asymmetry of the traveltime residuals, which introduce a varying signal-to-noise ratio within the data set. Thus, ray-density diagrams cannot be used to infer detailed confidence bounds on the velocity structure.

As an alternative to the statistical bounds discussed above, one can consider the bounds which result when point confidence bounds are placed on the data [constraints of equation (2)]. To specify that all the data must lie exactly within specified intervals is a very stringent requirement. The probability of this occurring is given by the product of probabilities that each datum lie within its confidence interval. For many data with Gaussian errors, wide confidence intervals must be placed on the data in order not to constrain the models unduly. In applying the technique to the Grimsel traveltime data, three standard errors were used as confidence intervals on the observations (0.06 ms). The probability that a given data point lies in this interval is 0.997. As might be expected, the resulting confidence bounds were quite large, with half-widths ranging from 0.4 km/s to 1.4 km/s. Perhaps the local bounds would be more useful for highly underdetermined problems with few data.

DISCUSSION AND CONCLUSIONS

A tomographic method has been presented for the computation of the range of velocities consistent with observed seismic traveltimes and their accompanying confidence bounds. The technique requires the same computer resources, memory, and disk space as the standard row-action methods for traveltime tomography. The amount of time required to compute each block bound is of the same order as that required for the determination of a velocity model using a row-action tomographic algorithm. The iterative algorithm is no more complex than row-action approaches to tomographic inversion and may be programmed without difficulty. Upper and lower bounds can be computed iteratively and may be used to estimate the uncertainties on velocity structure. The bounds derived compare favorably with those computed using the simplex algorithm for linear programming.

A variety of inversion algorithms were applied to crosshole data from the Grimsel mine experiment. The methods

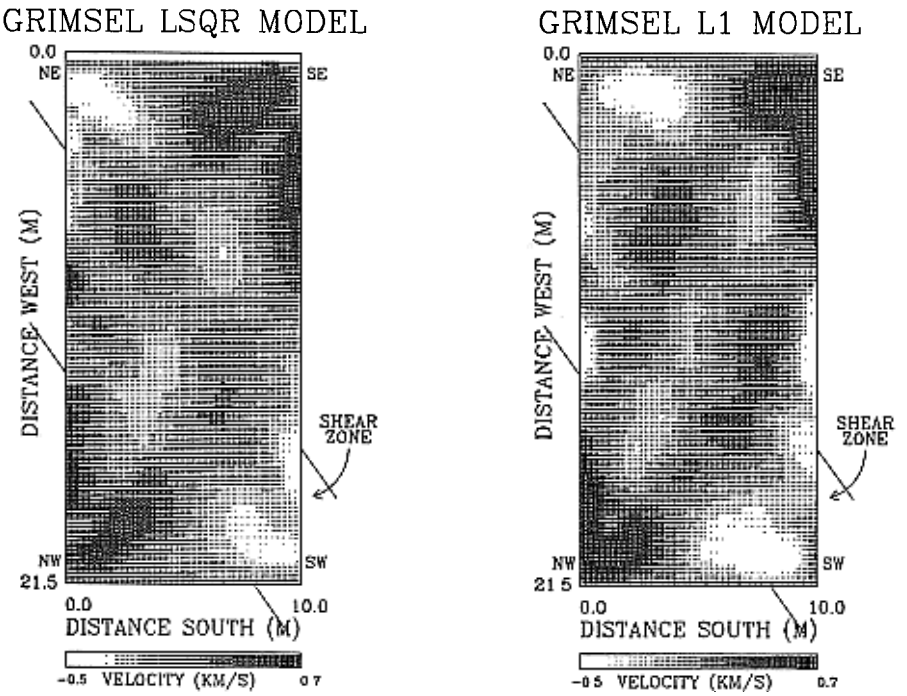

FIG. 8. Two solutions for the velocity perturbations in the Grimsel area: a conjugate gradient-based (LSQR) solution (left panel) and the velocity perturbations which minimize the ℓ^1 norm of the residual misfit (right panel). As in Figure 7, the mylonite shear zone is indicated on both panels.

produced varying images of the velocity structure due to inherent properties of each algorithm. The SVD and LSQR solutions require that the errors on the data are Gaussian; deviations from this will have a pronounced effect on the images. The ℓ^1 minimum residual norm solution is less sensitive to non-Gaussian errors, specifically outlying traveltime residuals, as caused by blunders in picking arrival times. To assess the solutions, upper and lower velocity bounds were computed for the subsurface distribution. The width of the bounds correlated well with the resolution and errors of the model parameters and with the ray density in the blocks. The bound means and half-widths indicate that the mylonitic zone is imaged in this experiment, with a range in velocity perturbations of 0.1 to 0.2 km/s possible. Significant positive and negative velocity perturbations appear in each alternate quadrant of the LSQR and minimum ℓ^1 residual norm solutions as well as in the mean statistical velocity bounds. An examination of the velocity perturbation bound widths (Figure 10) indicates that velocity perturbations in these quadrants are poorly constrained by the data. The symmetry of the velocity perturbation pattern there, identical to the calculated seismic anisotropy, indicates that some anisotropy is mapped into the solution. Specifically, velocity anisotropy is being mapped into the poorly resolved regions of these models. This points to the difficulty of separating lateral heterogeneity which contains inherent symmetry from seismic anisotropy due to factors such as crystal orientation and the need to invert for anisotropy as well as for lateral heterogeneity.

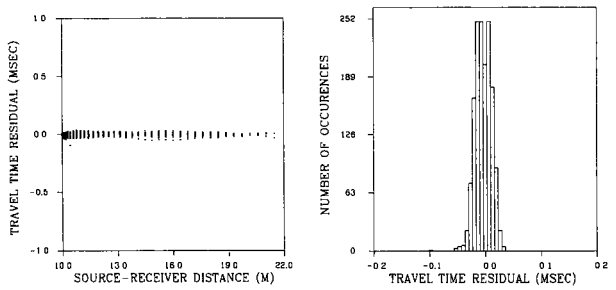

FIG. 9. (Left panel). Traveltime residuals computed relative to a back-ground velocity of 5.3 km/s with the LSQR model velocity perturbations added. (Right panel). Residual traveltime histogram when the velocity perturbations from the LSQR model (Figure 8) are added to the background velocity of 5.3 km/s and the result is used as a velocity model with which to compute traveltimes.

One parameter in the iterative scheme not discussed is the relaxation parameter s which appeared in the row-action algorithm [equation (6)]. Unfortunately, there is no general formula for estimating s other than the constraint that it lie between 0.5 and 2.0. To determine the effect of variations of s on the convergence of the algorithm, a single block (block 95, row 10, column 5) was examined. The lower bound for this block was computed using the formula in equation (6) with various values for s. The values of the bounds as a function of the iteration number for s equal to 0.75, 1.00, and 1.50 are shown in Figure 11. Clearly, the ultimate convergence does not depend strongly on the exact value of s. However, if properly chosen, the rate of convergence can be somewhat enhanced.

Velocity bounds may be computed for very large data sets such as the ISC traveltimes (Vasco et al., 1989). Such problems may involve hundreds of thousands to a few

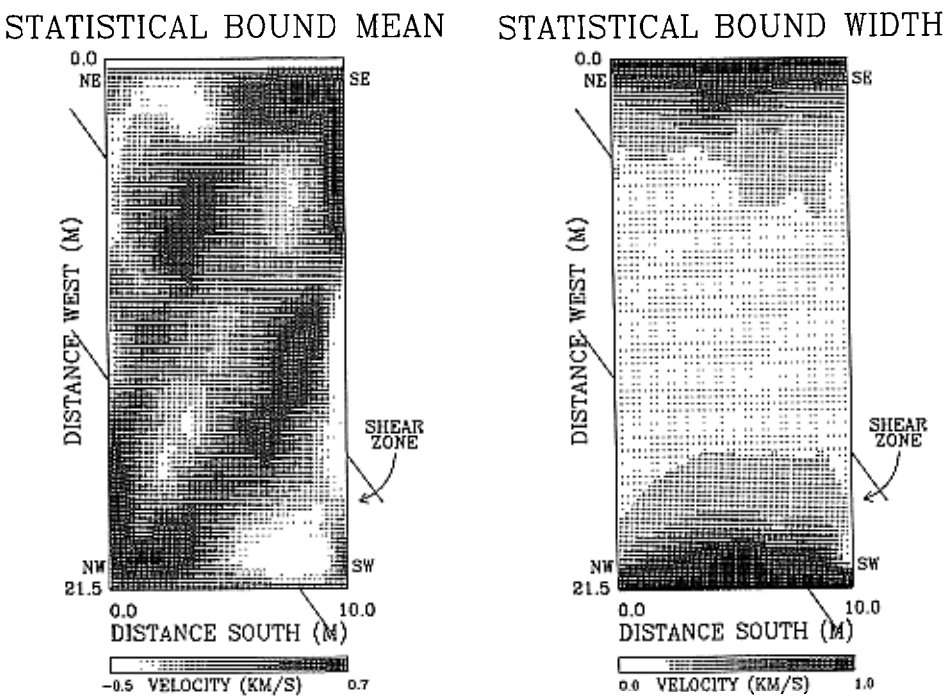

FIG. 10. (Left panel). Mean value of the statistical velocity bounds. (Right panel). Half-width of the statistical velocity bounds: (upper − lower)/2. The mylonite shear zone is denoted by the arrows.

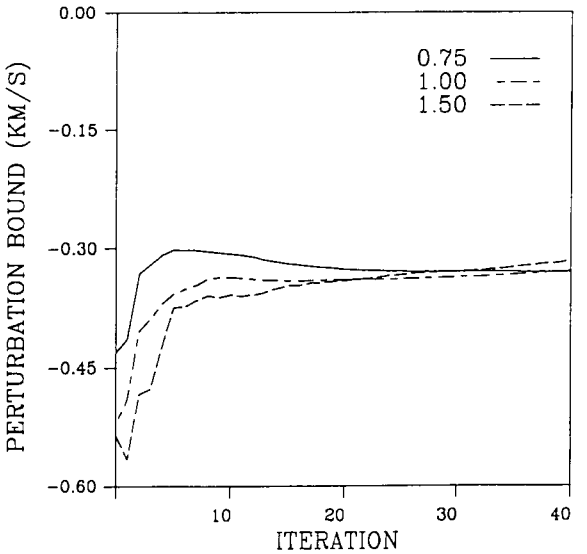

FIG. 11. Convergence of the algorithm for block 95 (row 10, column 5) for various relaxation parameter s values.

million raypaths, and hence it is not possible to compute the generalized inverse or the resolution and covariance matrices. The iterative computation of bounds may prove to be one of only a few ways to estimate uncertainties on large-scale tomographic solutions.

ACKNOWLEDGMENTS

I wish to thank Ernest Majer and John Peterson for supplying the very high-quality Grimsel traveltime data. This report was supported by the Assistant Secretary for Energy Research, Office of Basic Energy Sciences, and the Office of Civilian Radioactive Waste Management, Office of Geologic Repositories, U.S. Department of Energy, under contract no. DE-AC03-76SF00098. Center for Computational Seismology contribution number 95.

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology–Theory and methods: **I** and **II**: W. H. Freeman and Co.

Backus, G., 1965, Possible forms of seismic anisotropy in the uppermost mantle under oceans: J. Geophys. Res., **70**, 3429–3439.

Claerbout, J. F., and Muir, F., 1973, Robust modelling with erratic data: Geophysics, **38**, 826–844.

Clayton, R. W., and Comer, R. P., 1983, A tomographic analysis of mantle heterogeneities from body wave traveltime (abst): EOS, Trans. Am. Geophys. Union, **64**, 776.

Dantzig, G. B., 1963, Linear programming and extensions: Princeton Univ. Press.

Dines, K., and Lytle, J., 1979, Computerized geophysical tomography: Proc. Inst. Elec. Electron. Eng., **67**, 1065–1073.

Gill, P. E., Murray, W., and Wright, M. H., 1981, Practical optimization: Academic Press.

Herman, G. T., and Lent, A., 1978a, A family of iterative quadratic optimization algorithms for pairs of inequalities with application in diagnostic radiology: Math. Prog. Stud., **9**, 15–29.

———1978b, A relaxation method with application in diagnostic radiology, *in* Prekopa, A., Ed., Survey of mathematical programming, **III**: North-Holland.

Hildreth, C., 1957, A quadratic programming procedure: Naval Res. Logist. Quart., **4**, 79–85.

Lent, A., and Censor, Y., 1980, Extensions of Hildreth's row-action method for quadratic programming: SIAM J. Control and Optim., **18**, 444–454.

Majer, E. L., Peterson, J. E., Karasaki, K., Myer, L. R., Long, J. L., and Martel, S., 1990, Results of the fracture research investigation: LBL Report, Berkeley, CA.

Mangasarian, O. L., 1981, Iterative solution of linear programs: SIAM J. Number. Anal., **4**, 606–614.

Mangasarian, O. L., and Meyer, R. R., 1979, Nonlinear perturbation of linear programs: SIAM J. Control. and Optim., **17**, 745–752.

Menke, W., 1984, Geophysical data analysis: Discrete inverse theory: Academic Press.

Murtagh, B. A., and Saunders, M. A., 1978, Large-scale linearly constrained optimization: Math. Prog., **14**, 41–72.

Nolet, G., 1985, Solving or resolving inadequate and noisy tomographic systems: J. Comp. Physics, **61**, 463–482.

Paige, C. C., and Saunders, M. A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Trans. Math. Software, **8**, 43–71.

Parker, R. L., 1972, Inverse theory with grossly inadequate data: Geophys. J. Roy. Astr. Soc., **29**, 123–138.

——— 1974, Best bounds on density and depth from gravity data: Geophysics, **39**, 644–649.

——— 1975, The theory of ideal bodies for gravity interpretation: Geophys. J. R. Astr. Soc., **42**, 315–334.

Parker, R. L., and McNutt, M. K., 1980, Statistics for the one-norm misfit measure: J. Geophys. Res., **85**, 4429–4430.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1986, Numerical recipes: Cambridge Univ.

Sabatier, P. C., 1977a, Positivity constraints in linear inverse problems: I. General theory: Geophys. J. Roy. Astr. Soc., **48**, 415–441.

——— 1977b, Positivity constraints in linear inverse problems: II. applications: Geophys. J. Roy Astr. Soc., **48**, 443–459.

Scales, J. A., 1987, Tomographic inversion via the conjugate gradient method: Geophysics, **52**, 179–185.

Spakman, W., and Nolet, G., 1988, Imaging algorithms, accuracy and resolution in delay time tomography, *in* Nolet, G., Ed., Mathematical geophysics: D. Reidel Publishing Co., 155-187.

Vasco, D. W., 1986, Extremal inversion of traveltime residuals: Bull. Seis. Soc. Am., **76**, 1323–1345.

Vasco, D. W., Pulliam, R. J., and Johnson, L. R., 1989. Computing bounds on Mantle velocity structure from ISC travel times using a row-action method (abst): EOS, Trans. Am. Geophys. Union, **70**, 1213.